

# **Complexity in Chemistry, Biology and Ecology**

**Danail Bonchev and Dennis H. Rouvray, Editors**

**Springer, New York, 2004**

## **Content**

**Chapter 1. On the Complexity of Fullerenes and Nanotubes**, Milan Randić, Xiaofeng Guo, Dejan Plavšić and Alexandru T. Balaban

- 1.1. Introduction
- 1.2. On the Complexity of the Complexity Concept
- 1.3. Complexity and Branching
- 1.4. Complexity of Smaller Molecules
- 1.5. Augmented Valence as a Complexity Index
- 1.6. Complexity of Smaller Fullerenes
- 1.7. Comparison of Local Atomic Environments
- 1.8. The Role of Symmetry
- 1.9. Concluding Remarks on the Complexity of Fullerenes
- 1.10. On the Complexity of Carbon Nanotubes
- 1.11. References

**Chapter 2. Complexity and Self-Organization in Biological Development and Evolution**, Stuart A. Newman and Gabor Forgacs

- 2.1. Introduction: Complex Chemical Systems in Biological Development and Evolution
- 2.2. Dynamic Multistability and Cell Differentiation
- 2.3. Chemical Oscillations and Segmentation
- 2.4. Reaction-Diffusion Mechanisms and Embryo Pattern Formation
- 2.5. Evolution of Developmental Mechanisms
- 2.6. Conclusions
- 2.7. References

**Chapter 3. The Circle That Never Ends: Can Complexity Be Made Simple?** Donald C. Mikulecky

- 3.1. Introduction: The Nature of the Problem and Why It Has No Clear Solution
- 3.2. An Introduction to Relational Systems Theory
- 3.3. The Structure of Network Thermodynamics as Formalism
- 3.4. Simulation of Non-Linear Networks on Spice
- 3.5. Relational Networks and Beyond

### 3.6. References

## **Chapter 4. Graphs as Models of Large-Scale Biochemical Organization**, Pau Fernández and Ricard V. Solé

- 4.1. Introduction
- 4.2. Basic Properties of Random Graphs
- 4.3. Protein Structure and Contact Graphs
- 4.4. Protein Interaction Networks
- 4.5. Gene Networks
- 4.6. Overview
- 4.7. References

## **Chapter 5. Quantitative Measures of Network Complexity**, Danail Bonchev and Gregory A. Buck

- 5.1. Some History
- 5.2. Networks as Graphs
- 5.3. How to Measure Network Complexity
- 5.4. Combined Complexity Measures Based on the Graph Adjacency and Distance
- 5.5. Vertex Accessibility and Complexity of Directed Graphs
- 5.6. Complexity Estimates of Biological and Ecological Networks
- 5.7. Overview
- 5.8. References

## **Chapter 6. Cellular Automata Models of Complex Biochemical Systems**. Lemont B. Kier and Tarynn M. Witten

- 6.1. Reality, Systems, and Models
- 6.2. General Principles of Complexity
- 6.3. Modeling Emergence in Complex Biosystems
- 6.4. Examples of Cellular Automata Models
- 6.5. Summary
- 6.6. References

## **Chapter 7. The Complex Nature of Ecodynamics**, Robert E. Ulanowicz

- 7.1. Introduction
- 7.2. Measuring the Effects of Incorporated Constraints
- 7.3. Ecosystems and Contingency
- 7.4. Autocatalysis and Non-mechanical Behavior
- 7.5. Causality Reconsidered
- 7.6. Quantifying Constraint in Ecosystems
- 7.7. New Constraints to Help Focus a New Perspective
- 7.8. References

## Preface

As we were at pains to point out in the companion volume to this monograph, which was entitled “Complexity in Chemistry: Introduction and Fundamentals”, complexity is to be encountered just about everywhere. All that is needed for us to see it is a suitably trained eye and it then appears almost magically in all manner of guises. Because of its ubiquity, complexity has been and currently still is being defined in a number of different ways. Some of these definitions have led us to major and powerful new insights. Thus, even in the present monograph, the important distinction is drawn between the interpretations of the concepts of complexity and complication and this is shown to have a significant bearing on how systems are modeled. Having said this, however, we should not fail to mention that the broad consensus that now gained acceptance is that all of the definitions of complexity are in the last analysis to be understood in essentially intuitive terms. Such definitions will therefore always have a certain degree of fuzziness associated with them. But this latter desideratum should in no way be viewed as diminishing the great usefulness of the concept in any of the many scientific disciplines to which it can be applied. In the chapters that are included in this monograph the fact that differing concepts of complexity can be utilized in a variety of disciplines is made explicit. The specific disciplines that we embrace herein are chemistry, biochemistry, biology, and ecology.

The chapter titled “On the Complexity of Fullerenes and Nanotubes” is written by an international team of scientists led by Milan Randić. While devoted to specific chemical applications of complexity theory, and dealing with hot topics in contemporary chemical technology, the chapter complements contributions to the quantification of complexity made in the preceding volume “Complexity in Chemistry”. Most approaches to the notion of complexity in molecules and molecular graphs have been based on an evaluation of selected graph invariants, calculated for the graph itself or, most recently, for all of its subgraphs. This chapter focuses more on the influence that symmetry elements have on complexity of the objects considered. This is a controversial theme, with opposing opinions in the literature, because of the prevailing view on symmetry as a simplifying factor. The authors offer an improvement of symmetry-based complexity measures by accounting for the cardinality of the sets of equivalent elements. In addition, the concept of presenting the complexity measure as a complexity vector or sequence, originally developed for subgraph-based complexity measures, is now realized for distance-based sequences. The latter contain the average count of the number of nearest neighbors at various distances, and in the case of the fullerenes also the average distance between the twelve pentagonal faces of the fullerenes. The review ends with a discussion of the complexity of nanotubes, for which the main role appears to be played by the twist and counter-twist parameters that determine the nanotube helicity and diameter. As in the case of the fullerenes, the authors conclude that no single parameter seems to be sufficient to characterize nanotube complexity.

Newman and Forgacs focus in their chapter on the physicochemical aspects of complexity in developmental and evolutionary biology. The major emphasis in development studies has traditionally been on the hierarchical regulatory relationships among genes, while the variation of genes has played a corresponding role in evolutionary research. Recently, however, investigators have focused on the roles played by the physical and dynamical properties of cells and tissues in producing biological characteristics during ontogeny and phylogeny. The interactions among gene products, metabolites, ions, etc., reaction-diffusion coupling, and

opportunities for molecular diffusion over macroscopic distances, lead to self-organizing multistable, oscillatory, and pattern forming dynamics. These system properties, not specified in any genetic program, can account for most of the features of animal body structures, including cell differentiation, tissue multilayering, segmentation, and left-right asymmetry. The authors point out that these chemical-dynamic properties are *generic* and are common to living and nonliving systems. As such, they have played a major role in the evolution of ancestral multicellular organisms, more than for modern organisms with their hierarchical genetic control. The morphologies originally generated by physicochemical dynamics probably provided morphological templates for genetic evolution that stabilized and reinforced (rather than innovated) multicellular body plans and organ forms. The hierarchical genetic control of development seen in modern organisms can thus be considered as an outcome of this evolutionary interplay between genetic change and generic physicochemical processes.

The chapter “The circle that never ends: Can complexity be made simple?” by Mikulecky begins with the premise that all real systems are complex and that there can be no single approach to such systems. Mikulecky presents an introduction to a relatively new approach called “relational systems theory” to which he has made valuable contributions along with such pioneers as Katchalsky, Peusner, and Rosen. As an essential part of the efforts being made to create a new, non-Newtonian paradigm of science, this theory accounts for the irreducibility of certain context dependent functional components in the system, components that would disappear when the context defining them is destroyed by reductionist techniques. The influence of reductionism on the methods of science is described by Mikulecky in some detail in an effort to provide reasons for moving beyond the restrictions to systems imposed on us by this traditional approach. The relational approach clearly identifies the nature of the *functional* components as that entity which makes a complex whole more than the sum of its parts. The relational model is expressed in terms of processes rather than its physical parts. There is no way to uniquely identify a functional component by the way it relates to the physical parts of the system even though a relationship has to exist. Relational models that are context dependent and self-referential are considered inherently incapable of being reduced to the algorithmic procedures that make mechanistic systems so adaptable to computer simulation and other computational techniques. This non-computability, which makes it impossible to explain fully a system proceeding from a single model, is thus regarded as a key characteristic of complex reality.

The study of the organization extant in mechanisms holds out the promise of bridging the gap between the mechanistic approach characteristic of reductionism and the new relational approach. For that reason, Mikulecky makes use of Network Thermodynamics to illustrate the relation between organization and the material parts in mechanisms as a way of showing that the material parts alone are insufficient to give a system description even for mechanisms. Network Thermodynamics combines topology with analytical mathematics to model large complicated systems in a way that demonstrates the role of organization in dynamic models. Versions of Network Thermodynamics have been developed during the last 30 years by Oster, Perelson and Katchalsky, Peusner, and Mikulecky. Summarizing these results, this review focuses on a generalized formalism of electrical circuit theory, which is shown to be applicable to all networks representing physical systems. The reader can thus gain new and fruitful insights into the possibilities for modeling complex dynamic networks. One can only agree with Mikulecky’s

conclusion that “The challenge arising from acknowledging the complexity of the real world is to try to maintain scientific rigor without stripping the investigative process of the tools needed to deal with context dependence and self-reference. In that way, the circle never ends.”

Complex biochemical systems are difficult to study in their entirety due to the overwhelming number of constituents they have. However, in focusing on the interactions between the constituents one arrives at the underlying networks, the best representation of the integrated whole. This is the starting point in the chapter of Fernandez and Solé devoted to networks (graphs) as models of large-scale biochemical organization. With many examples, from the contact network in a folded protein to protein-protein interaction networks to gene regulatory networks, it is convincingly demonstrated how some of the properties of complex systems can be approached through the study of the properties of their corresponding networks. One of the very few reviews on the topological properties of complex networks, this article begins by introducing the necessary notions of graph theory. Following the chronology of development of complex network theory, the authors define first the properties of random graphs. Vertex degree distribution is analyzed in detail, including clustering as a basic property of complex graphs. The transition from random graphs to apparently nonrandom, real-life networks is elegantly presented beginning with the random “long-distance shortcuts”, and then going on to Strogatz and Watts’ “small world” theory. The folded protein structure is the first example of a complex biochemical network the authors describe. The linear chain of residues is folded and this final shape involves contact, long-range interactions among the aminoacid residues. The mandatory small-world, connectivity and clustering analysis is followed by that of hierarchical clustering, which reveals the hidden modular structure of the network. Protein-protein networks follow with examples of such network fragments in humans and *Saccharomyces cerevisiae*. Here, the latest techniques of network assortativeness and correlation profiles are introduced. The proteome emergence and evolution is modeled proceeding from an analogy with the gene duplication mechanism in genetic evolution. The last section covers the gene regulatory networks that determine the functioning of the living cell. These networks are analyzed as Kauffman’s Boolean networks and possess a number of dynamic properties such as the emergence of high-order, robustness, and phase transitions. Three distinct phases are shown to exist in such networks: ordered, chaotic, and critical and the conditions for their existence are mathematically formulated. The formalism is illustrated with the regulatory networks of *E. coli* and *S. cerevisiae*. Using networks as a theoretical framework for complex biochemical systems is relevant for a number of reasons, and especially for providing well-defined quantitative properties to be reproduced by dynamical models of network evolution, thereby serving as a blueprint for the laws of organization of living matter.

Complexity of molecules and the criteria for quantitative complexity measures have been discussed in detail in the preceding volume “Complexity in Chemistry”. By this reason, the chapter “How to Measure Complexity?” by Bonchev and Buck only briefly reviews the history of the topic. The emphasis is put on the latest development that identified nonrandom dynamic networks as a universal language to describe complexity of systems as diverse as the discrete space-time, molecules, living matter, ecosystems, social systems, financial markets, and World Wide Web. This revolutionary idea, advanced by Barabási in 1999, had a profound impact on life sciences and first of all on mathematical biology, which is more and more dominated by systems biology and bioinformatics. This has also changed the approaches used to quantify

complexity in biology and ecology. In addition to information theory, traditionally used to define complexity as *compositional* diversity, graph theory emerged as a universal tool for assessing *structural* or *topological* complexity of any dynamic evolutionary system. The authors proceed with analysis of the role of symmetry as simplifying factor, diminishing systems complexity. They warn against the blind use of information theory for complexity estimates, and advocate the use of the information on the vertex degree distribution as one of the very few satisfactory information measures of topological complexity. A general scheme is presented for quantifying complexity as global, average, and normalized complexity indices. Three graph theoretical measures of network complexity are recommended: the subgraph count, the overall connectivity index, and the walk count. All three are presented both as a global index, and as an ordered sequence of terms related to subgraphs, and respectively walks, from the smallest to the largest size. This makes it possible to assess networks complexity by the first several terms of each of these indices, avoiding thus the combinatorial explosion for large-scale networks. Two complexity measures combining graph adjacency and graph distance (graph radius) are also presented for the first time. These indices unite the intuitive ideas of structural complexity resulting from high connectivity and small vertex separation (the “small world” concept). Complexity of directed networks is analyzed by introducing a measure for vertex accessibility, which enables producing realistic total distance and connectedness estimates of these networks. The mathematical formalism introduced throughout the chapter is illustrated in detail by examples of protein-protein networks and food webs.

The chapter titled “Cellular Automata Models of Complex Biochemical Systems” by Kier and Witten begins by introducing the basic notions of systems, defining the states of the system, the system observables and their interactions. The basic features of modeling and simulation are discussed. Focusing on modeling in chemistry and molecular biology, two powerful methods for chemical investigations are discussed and compared: molecular dynamic and Monte Carlo simulations. Both approaches have great strengths and often lead to quite similar results for the properties of the systems studied. However, these methods depend on rather elaborate models of the molecular interactions. As a result, both methods are highly demanding computationally, and research-level calculations are normally run on supercomputers, clusters, or other large systems. Before introducing the alternative approach of *cellular automata* that greatly simplifies the view of the molecular system, and significantly reduces the computational demand, the authors first address the general principles of complexity. The notion of a complex system is distinguished from that of a complicated system, which is no more than the sum of its parts. In contrast, in complex systems “the whole is greater than the sum of the parts.” What is more these systems possess the properties of *emergence*, *adaptation*, and *self-organization*. The authors discuss these properties of complex systems along with the existing hierarchy of complexity focusing on the structure and interconnectedness of the “layers” of complexity. The contributions of the pioneers of complexity theory are elucidated and illustrated by numerous examples.

This chapter continues by introducing the basic notions and principles of cellular automata. Cellular automata consist of a discrete lattice of cells and evolve in discrete time steps. Each site takes on a finite number of possible values. The value of each site evolves according to the same simple, heuristic rules, which depend only on a local neighborhood of sites around it. The results of a CA model are new sets of states of the constituents called the configuration of the system. This configuration arises from many changes and encounters among the constituents of the CA,

which may occur over a very long period of “time” in the model. The last part of the chapter presents abundant examples of applications of cellular automata to chemistry and biology, taken from Kier’s laboratory. The early work was directed toward the study of water and solution phenomena. Studies include cellular automata models of water as a solvent, dissolution of a solute, solution phenomena, the hydrophobic effect, oil and water de-mixing, solute partitioning between two immiscible solvents, micelle formation, diffusion in water, membrane permeability, acid dissociation, and dynamic percolation. Later work involved cellular automata models of molecular bond interactions, diffusion in water, including drug molecule diffusion and the hydrophobic effect, as well as generalizations of Kier’s chreode theory of diffusion in water and his theory of volatile anesthetic action. The chapter ends with a demonstration of the full potential of this powerful technique for application to complex biochemical pathways.

Over the last three centuries Newtonian dynamics has strongly guided how we look at nature. Newtonian systems are explained in terms of mechanistic or material causes only. They are deterministic, reversible in time, decomposable into their parts and composable again. Physical laws are assumed to apply everywhere, at all times and over all scales. Ulanowicz shows in his chapter that none of these notions applies to ecodynamics. His conclusion stems from the fact that constraints in living systems are not rigidly mechanical in nature, owing mainly to the cyclical relationships among some of them. Living systems are not fully constrained, i.e., they retain sufficient flexibility to adapt to changing circumstances. Flexibility is probably easier to discern in ecosystems than in organisms where the constraints are more prevalent and rigid. Autocatalysis plays an essential role in the emergence of non-mechanical behavior in all living systems. Whenever two or more autocatalytic loops arise from the same pool of resources, autocatalysis induces competition and symmetry-breaking. Ulanowicz asserts that the full ontogenetic mapping from genome to phenome is very likely a chimera. What is really needed is try to discover the principles of self-organization. The author shows that the imbalances in material and energy demanded by physics usually equilibrate at rates that are much faster than changes occurring in the constraints that actually determine the pattern of system exchanges. Thus, the pathway to achieving a mathematical description of ecodynamics lies in the quantification of internal constraints. The control of ecodynamics appears to be relational in nature -- how much any change in one constraint affects others with which it is linked. While it is impossible to treat explicitly all the constraints hidden within an ecosystem, their overall effects on the network of trophic interactions can nevertheless be quantified in a manner similar to that used in thermodynamics. For this purpose, Ulanowicz uses information theory arguments to introduce the concept of *ascendance* as a quantitative measure of how efficiently and coherently the system processes its medium. A phenomenological principle is proposed that “in the absence of major perturbations, ecosystems have a propensity to increase in ascendancy.” The ensuing description of ecodynamics, however, does not accord with the normal assumptions in Newtonian metaphysics, and Ulanowicz offers a reformulated ecological metaphysics in its place.

In conclusion, we would like to express the hope that the chapters contained within our monograph will not only make for some stimulating reading but that they will also afford our readers with valuable resource material for future research endeavors. We take this opportunity to thank all of our contributors for their valuable contributions to the fascinating subject of

complexity. We can only hope that our readers will derive much pleasure in perusing the exciting offerings herein.

Danail Bonchev

Dennis H. Rouvray